MAI 23770 — Ch. 1

Discrete vs. Continuous, dependent vs. independent

Multiplication Rule: if 2 exp, first has n_1 outcomes, second has n_2 outcomes, then first followed by second has $n_1 n_2$ outcomes. Same for r outcomes: $n_1 n_2 \cdots n_r$

Permutations: Order Matters!

- # of permutations of n distinct objects = $n!$
- # perm of n objects, n_1 are alike, n_2 alike (but different from others, [...] n_r = $\frac{n!}{n_1! n_2! \cdots n_r!}$

ex) aa b c c c ($n=6$, n_1, n_2, n_3) $\Rightarrow \frac{6!}{2!1!3!} = \frac{720}{12} = 60$

- # perm n distinct objects taken r at a time is: $_nP_r = \frac{n!}{(n-r)!}$

ex) # perm a,b,c 2 at a time = $\frac{3!}{(3-2)!} = 6$

Combinations: Order does not matter!

- # comb n objects taken r at a time = $\binom{n}{r} = \frac{n!}{r!(n-r)!}$

Probability

- $P(A) \geq 0 \ \forall A \in S$
- $P(S) = 1$
- if $A_1, A_2, \ldots, A_n$ are mutually exclusive, $\Rightarrow P(A_1 \cup A_2 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + \ldots + P(A_n)$
- ★ if S has N equally likely outcomes, A contains n outcomes, $P(A) = \frac{n}{N}$ ← use perm & comb to find n & N

Additive Rules

- $P(A') = 1 - P(A)$
- $P(\emptyset) = 0$
- $A_1 \subseteq A_2 \Rightarrow P(A_1) \leq P(A_2)$
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

Conditional Probability

$P(B|A) = \frac{P(A \cap B)}{P(A)}$, $P(A) > 0$

$P(A \cap B) = P(B|A)P(A)$, $P(A) > 0$

$P(A'|B) = 1 - P(A|B)$

- $P(\emptyset|B) = 0$
- $A_1 \subseteq A_2 \Rightarrow P(A_1|B) \leq P(A_2|B)$
- $P(A \cup B|C) = P(A|C) + P(B|C) - P(A \cap B|C)$

Notation

- S = Sample Space = set of all possible outcomes
- $A, B, \ldots$ = Event = one outcome or group of outcomes
- $P(A)$ = Probability of event A
- R_X = Set of all possible values of X
- X = Random variable
- x = one of the values of a random variable
- $P(X, Y) = P(X \cap Y)$
- μ = Mean = Expected value

* $E[(X)]$ is also denoted $E(X)$

Product Rule

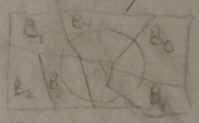
$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|(A_1 \cap A_2)) \ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$

if independent: $P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2) \ldots P(A_n)$

Bayes' Rule

if $B_1, \ldots, B_n$ are disjoint & $B_1 \cup B_2 \cup \ldots \cup B_n = S$ then,

$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$



ex) 3R, 2G, no replacement, what is P(A), where A is prob 2nd ball is red whereas P(B) is prob 1st is red

$P(A) = P(A|B)P(B) + P(A|B')P(B') = (\frac{2}{4})(\frac{3}{5}) + (\frac{3}{4})(\frac{2}{5}) = \frac{12}{20}$

Bayes's Theorem (if given same scenario as above)

$P(B_r|A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{k} P(B_i)P(A|B_i)}$

ex) $P(A) = 0.3$, $P(B) = 0.5$, $P(C) = 0.2$. If C is scratched what is $P(B)$?

$P(B|A \cup B) = \frac{P(B)}{P(A \cup B)} = \frac{0.5}{0.5 + 0.3} = \frac{5}{8}$

Independent $\Leftrightarrow P(B|A) = P(B)$ (same as $P(A|B) = P(A)$)

$\Leftrightarrow P(A \cap B) = P(A)P(B)$

if independent, $P(A \cap B) = P(A)P(B)$

see next page for $P(A \cap$

MAT 2377 (Ch 2)

A random variable associates a $\mathbb{R}$ # w/ each ϵ in S.

DISCRETE:

Probability Mass Function (PMF) (AKA prob. distribution)

PMF of X is $f(x) = P(X = x)$ $\forall$ possible values of x

- $f(x) \geq 0$
- $\sum f(x) = 1$
- $P(X = x) = f(x)$

Cumulative Distribution Function (CDF)

NOTE: $F(x) \neq f(x)$

CDF of X is $F(x) = P(X \leq x)$ $\forall$ values of x

- $F(x) = \sum_{t \leq x} f(t)$ $\quad -\infty < x < \infty$

CONTINUOUS: $P(a < X < b)$

$P(X = a) = 0$ for cont.

Probability Density Function (PDF)

$f(x)$ is PDF of continuous rand. var. X if

1) $f(x) \geq 0 \quad \forall \mathbb{R}x$
2) $\int_{-\infty}^{\infty} f(x)\,dx = 1$
3) $P(a < X < b) = \int_a^b f(t)\,dt$
4) CDF: $F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$

(Discrete) Joint Probability Distribution = PMF of X, Y if

1) $f(x,y) \geq 0 \quad \forall x, y$
2) $\sum_x \sum_y f(x,y) = 1$
3) $P(X = x, Y = y) = f(x,y)$
4) $P((X,Y) \in A) = \sum\sum_{(x,y) \in A} f(x,y)$ for any $A \in xy$ plane

(Continuous) Joint Probability Distribution of X, Y if

1) $f(x,y) \geq 0 \quad \forall x, y$
2) $\int\int f(x,y)\,dx\,dy = 1$
3) $P((X,Y) \in A) = \iint_A f(x,y)\,dx\,dy$ For any $A \in xy$ plane

Retrieving Individual Distributions from Joint Distributions, AKA "Marginal Distribution"

Discrete: $g(x) = \sum_y f(x,y)$, $h(y) = \sum_x f(x,y)$

(x ↑, y →; $g(x)$ ← marginal)

Continuous: $g(x) = \int f(x,y)\,dy$, $h(y) = \int f(x,y)\,dx$

x \ y	0	1	2	$g_X(x)$
0	1/35	8/35	6/35	15/35
1	10/35	20/35	0	30/35
2	10/35	0	0	10/35
marginal $h_Y(y)$	21/35	28/35	6/35	total 1

Conditional Distributions

- $f(y|x) = \dfrac{f(x,y)}{g(x)}$, $g(x) > 0$

similarly

- $f(x|y) = \dfrac{f(x,y)}{h(y)}$, $h(y) > 0$

Independent $\iff f(x,y) = g(x)h(y)$ $\forall (x,y)$

$\iff f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) \ldots f_n(x_n)$

Mean AKA Expected Value of a f(n) L of X: $L(X)$

$\mu = E[L(X)] = \begin{cases} \sum L(x) f(x) & \text{for } X \text{ discrete} \\ \int L(x) f(x)\,dx & \text{for } X \text{ continuous} \end{cases}$

($E(X)$; $L(x) = x$, $f(x) = P(x)$)

$x_{min} < \mu < x_{max}$

$\mu = E[L(X,Y)] = \begin{cases} \sum\sum L(x,y) f(x,y) & \leftarrow X, Y \text{ discrete} \\ \int\int L(x,y) f(x,y)\,dx\,dy & \leftarrow X, Y \text{ continuous} \end{cases}$

($E(X,Y)$)

Covariance of two rand. var. X, Y

$\sigma_{XY} = E[(X - \mu_X)(Y - \mu_Y)] = E(XY) - \mu_X \mu_Y$

Varience of X

$\sigma_X^2 = E[(X - \mu_X)^2] = E(X)^2 - \mu_X^2 = E(X)^2 - E(X^2)$

$\mu_X^2 = E(X^2) = \sum_x (x - \mu)^2 P(x)$

Correlation between X and Y is

$\rho = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$\rho > 0 \Rightarrow X \uparrow Y \uparrow$

$\rho < 0 \Rightarrow X \uparrow Y \downarrow$

$E(X^2) = \sum_{\forall x} x^2 p(x) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$

mean is sort of a weighted average of what x we should expect

Theorem Let X, Y be rand. var., a, b, c be constants, $h_1(x), h_2(x), h_3(x,y), h_4(x,y)$ be $\mathbb{R}$ f(n)s

- $E(aX + b) = aE(X) + b$
- $E[h_1(X) \pm h_2(X)] = E[h_1(X)] \pm E[h_2(X)]$
- $E[h_3(X,Y) \pm h_4(X,Y)] = E[h_3(X,Y)] \pm E[h_4(X,Y)]$
- $\sigma^2_{aX+bY+c} = a^2\sigma_X^2 + 2ab\sigma_{XY} + b^2\sigma_Y^2$

Theorem Let X, Y be 2 independent rand. var., a, b constants

- $E(XY) = E(X)E(Y)$
- $\sigma_{XY} = 0$
- $\sigma^2_{aX+bY} = a^2\sigma_X^2 + b^2\sigma_Y^2$

generalized

- $E[\sum a_i X_i] = \sum a_i E[X_i]$ ($X_1 \ldots X_n$ not necessarily independent)
- if $X_1 \ldots X_n$ is independent, $\hookrightarrow \sigma^2_{\sum a_i X_i} = \sum a_i^2 \sigma^2_{X_i}$

MAT 2377 D (Ch 3) $f(x) = P(X = x)$ for "less than", use tables (Ch 3)

Bernoulli Trial: p = "success", q = "Failure" $= (1-p)$, $f(x) = p^x(1-p)^{1-x}$, $x = 0, 1$, $\mu = p$, $\sigma^2 = p(1-p)$

Binomial Distribution: Same as Bernoulli, but done n times and X counts # of successes

↳ $f(x) = \binom{n}{x} p^x (1-p)^{n-x}$, $x = 0, 1, \cdots, n$, $\mu = n \cdot p$, $\sigma^2 = np(1-p)$

Multinomial Distribution: Same as Binomial, but w/ k outcomes w/ probabilities $p_1, p_2, \cdots$

↳ $f(x_1, x_2, \cdots x_k) = \left(\frac{n!}{x_1!, \cdots, x_k!}\right) p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k}$, Covariance $(X_1, X_2) = \sigma_{x_1 x_2} = -n p_1 p_2$

Hypergeometric Distribution: Same as Binomial, but small population* & Select w/o replacement

n = # of items picked from box, N = # items in box, k = items of 1 kind ("success"), $(N-k)$ = items of other kind

x = # of items selected that are of type k $(x \le k)$, $n - x \le N - k$

↳ $h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$, $\mu = \frac{nk}{N}$, $\sigma^2 = \left(\frac{N-n}{N-1}\right) n \frac{k}{N}\left(1 - \frac{k}{N}\right)$

*for $\frac{n}{N} \le 0.05$, we can use binomial approximation

Geometric Distribution: "How many times until first success?" "What is prob of first success @ n^{th} trial"

constant probability of "success"

↳ $f(x) = p(1-p)^{x-1}$, $x = 1, 2, \cdots$ $\mu = p^{-1}$, $\sigma^2 = \frac{1-p}{p^2}$

Negative Binomial: for the k^{th} success: $f(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}$ $x \ge k$, $\mu = \frac{k}{p}$, $\sigma^2 = k\left(\frac{1-p}{p^2}\right)$

$P(X > x) = \sum_{x=x+1}^{\infty} p(1-p)^{x-1} = (1-p)^x = 1 - P(X \le x)$

$P(X \le x) = \sum_0^x P(X = x)$ ⇒ $P(X < x) = 1 - P(X > x) = 1 - (1-p)^x$

Poisson Distribution: x = # successes, λ = average of occurance $\left(\frac{\text{Avg occurances}}{\text{time}}, \text{time}\right)$ (in time t), t = time interval

$\frac{t}{n}$ = time segment. As $n \to \infty$, $p \to 0$, binomial approximates to poisson, & $n \cdot p$ remains constant

↳ $b(x; n, p) \to P(x; \mu) = \frac{e^{-\lambda} \lambda^x}{x!}$ $x = 0, 1, \cdots$ $\mu = \lambda t$, $\sigma^2 = \lambda t$

$P = \lambda\left(\frac{t}{n}\right)$

★ Can use poisson estimation if $n \ge 20$ & $p \le 0.05$ or if $n \ge 100$ & $p \le 0.10$

↑ DISCRETE ↑

↓ CONTINUOUS ↓ (sort of)

Continuous Uniform Distribution: interval $[A, B]$ $f(x; A, B) = \begin{cases} \frac{1}{B-A} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$

x = # success

$\mu = \frac{A+B}{2}$ $\sigma^2 = \frac{(B-A)^2}{12}$

Normal Distribution: $f(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)}$ ← Not needed. Just change to standard Normal & use table.

STANDARD NORMAL: $\mu = 0$, $\sigma = 1$ ↑ $Z = \frac{X - \mu}{\sigma}$ ← cdf = Φ

$P(x_1 < X < x_2) = \Phi\left(\frac{x_2 - \mu}{\sigma}\right) - \Phi\left(\frac{x_1 - \mu}{\sigma}\right) = P\left(z < \frac{x_2 - \mu}{\sigma}\right) - P\left(z < \frac{x_1 - \mu}{\sigma}\right)$

↳ $X \sim \sigma^2$ $P(X < x) = P\left(Z < \frac{x - \mu}{\sigma}\right)$

Normal Approximation to Binomial: if X is binomial rand.var. w/ $\mu = np$, $\sigma^2 = np(1-p)$, then as $n \to$

$q = (1-p)$ $Z = \frac{X - np}{\sqrt{npq}}$ ⇒ $P(x_1 < X < x_2) \cong \Phi\left(\frac{x_2 - np}{\sqrt{npq_n}}\right) - \Phi\left(\frac{x_1 - np}{\sqrt{npq_n}}\right)$

more on next page

MAT 2377 D (Ch 4)

Exponential Distribution: (based on poisson) $P(T>t) = \left(1-\frac{\lambda t}{n}\right)^n \to \boxed{e^{-\lambda t}}$

So $F(t) = P(T \le t) = 1 - e^{-\lambda t}$ & $f(x) = \lambda e^{-\lambda x}$, $\boxed{\mu = \lambda^{-1}}$, $\boxed{\sigma^2 = \lambda^{-2}}$

↳ **Memoryless Property:** (for exponential) $P(T \ge t_0 + t \mid T \ge t_0) = e^{-\lambda t} = P(T \ge t)$ "Probability that it survives t time given it has already survived to is same as probability it will survive t time!"

Gamma Distribution: rand.var. X has gamma distr. w/ param. α & β if density given by

↳ gamma f(): $\boxed{\Gamma(\alpha) = (\alpha-1)!} = \int_0^\infty x^{\alpha-1} e^{-x}\,dx$ $\boxed{\mu = \alpha\beta}$, $\boxed{\sigma^2 = \alpha\beta^2}$ → $f(x;\alpha,\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}$

properties of factorials → Note: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(n+1) = n!$

α = # of poisson events, β = mean time between failures

Exponential is gamma if $\alpha = 1$, $\beta = \frac{1}{\lambda} \Rightarrow \mu = \beta = \lambda^{-1}$, $\sigma^2 = \beta^2 = \lambda^{-2}$

Chi-Squared Distribution: chi-squared is a special case of gamma w/ $\alpha = \frac{\nu}{2}$ (ν = deg. of freedom = $n-1$), $\beta = 2$

↳ $\boxed{\mu = \nu}$, $\boxed{\sigma^2 = 2\nu}$ if $X \sim N(\mu, \sigma^2)$, then $\left(\frac{X-\mu}{\sigma}\right)^2 \sim \chi_1^2$

Sample Mean: $\boxed{\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i}$ (same as $\bar{x}$) — changes w/ sample, rand.var

Sample Variance: $\boxed{S^2 = \frac{1}{n(n-1)}\left[n\sum X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2\right]}$ (Ch 4)

$\boxed{S^2 = \frac{\sum_{i=1}^{n}[(X_i - \bar{X})^2]}{(n-1)}}$ (like σ^2, but divide by $n-1$ instead of n)

Central Limit Theorem: $X_1, \ldots, X_n$ is random sample from population w/ μ (Pop.mean) & σ^2 (Pop.var)

↳ When $n \ge 30$, $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$ note: $\frac{\sigma}{\sqrt{n}} = \sigma_{\bar{x}}$ (σ of pop., σ of the mean)

$\boxed{Z = \left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}\right)}$ ↑ standard normal

↳ **Variance of S.M.:** $\boxed{Var(\bar{x}) = \frac{\sigma^2}{n}}$ (Pop.var)

properties of $E(x)$, $V(x)$: $\boxed{E(ax+b) = a\cdot E(x) + b}$, $\boxed{V(ax+b) = a^2 V(x)}$

Central Limit Theorem (different populations): $n_1 \ge 30$, $n_2 \ge 30$

$$\boxed{Z = \left(\frac{(\bar{X}_{n_1} - \bar{X}_{n_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right)}$$

(use standard normal)

Sampling Distribution of S^2 (sample variance)

↳ $S^2 \Rightarrow \chi^2_{n-1} \sim \boxed{\frac{(n-1)S^2}{\sigma^2}} = \boxed{\frac{\sum(X_i - \bar{X}_n)^2}{\sigma^2}}$ ($\nu = n-1$)

T-Distribution: Z = standard normal rand.var., V = Chi-Squared rand.var., Z, V independent

↳ $\boxed{T = \frac{Z}{\sqrt{\frac{V}{\nu}}}}$ (ν = deg. of freedom)

$h(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}}\left(1 + \frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}$

$\boxed{T = \left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}\right)}$

F-Distribution: U, V = 2 (independent) rand.var w/ chi-squared distributions, w/ ν_1, ν_2 respectively

↳ $\boxed{F = \frac{(U/\nu_1)}{(V/\nu_2)}}$

note: if S_1^2 & S_2^2 are sample variances w/ sizes n_1, n_2, σ_1^2 & σ_2^2, $\Rightarrow \boxed{F = \frac{(S_1^2/\sigma_1^2)}{(S_2^2/\sigma_2^2)}}$ $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$

MAT 2377 D (Ch5) estimation ⇒ confidence interval ∧ means it was approximated

Notation: Random variable = capital letter, lowercase = one instance of the rand. var
↳ eg. Sample Mean $\bar{X}$, one value of $\bar{X}$ is $\bar{x}$. For sample size n, we may subscript n → Eg. $\bar{X}_n, S_n^2, \bar{x}_n$

θ, μ, σ^2 → Parameters: "fixed", represented w/ greek letters, estimates = capital, letter = specific value
↳ eg. θ = parameter, estimate = $\hat{\Theta}$ (biased) ← rand. var (lowercase), specific value = $\hat{\theta}$, estimate is based on observed data sample $x_1, \dots, x_n$
↳ eg. p = population proportion, point estimator $\hat{P}$ (unbiased), specific observed value = $\hat{p}$

Unbiased Estimators: $\hat{\Theta}$ is unbiased estimator of θ if $E(\hat{\Theta}) = \theta$, ← unbiased estimator w/ smallest σ^2 is called "most efficient estimator" of θ
↳ $\bar{X}_n$ is unbiased estimator for pop. mean μ, S^2 is unbiased estimator for pop. var σ^2

if 0 is in it, reject

↳ Confidence Intervals (CI): $P(\hat{\theta}_L < \theta < \theta_R) = 1-\alpha$ is a $100(1-\alpha)\%$ confidence interval. (want small interval, high degree of confidence.)
↳ $\theta \pm \theta_R - \theta$

Estimating Population Mean $\bar{X}_n$: (Using $\bar{x}_n$ to estimate μ) @ $\bar{x}_n$ $\hat{\theta}_L$ is $\bar{x}_n - Z\dots$, $\hat{\theta}_R$ is $\bar{x}_n + Z\dots$

→ When variance (σ^2) is known: μ is included in interval $\bar{X}_n \pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ w/ probability $(1-\alpha)$
↳ error ($= \frac{\text{width}}{2}$) will not exceed e when the sample size is $n = \left(\frac{Z_{\alpha/2}\sigma}{e}\right)^2$, $n \in \mathbb{N}$
↳ standard error of $\bar{X}$, denoted $s.e.(\bar{X}) = \frac{\sigma}{\sqrt{n}}$

Finding $Z_{\alpha/2}$: 1- do $A = 1-(\alpha/2)$, 2- find A in "middle" of table (not edges), 3- Z = edge Vertical + Horizontal @ A
Finding probability: $P(Z < \alpha)$: 1- Plug Z in "edges" of table, 2- P = where Vertical & Horizontal meet.

→ When variance (σ^2) is NOT known: use T-distribution of the rand. var: $T = \frac{\sqrt{n}(\bar{X}_n - \mu)}{S}$
w/ $n-1$ deg. of freedom, μ is included in interval: $\bar{X}_n \pm t_{\alpha/2}\frac{S}{\sqrt{n}}$

Estimating difference between 2 (pop.) means $\bar{X}_1$ & $\bar{X}_2$:

Variances	Interval	Definition	d.f = ν
σ_1^2, σ_2^2 known	$\bar{X}_1 - \bar{X}_2 \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$		
$\sigma_1^2 = \sigma_2^2 = \sigma^2$ unknown	$\bar{X}_1 - \bar{X}_2 \pm t_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$S_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$	n_1+n_1-2
σ_1^2, σ_2^2 unknown	$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$		* OR **

$$* = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{(n_1-1)} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{(n_2-1)}}$$

** = min[(n_1-1) & (n_2-1)]

Paired Observations: Since X and Y are not independent, use their difference (D = X − Y) as a random variable. eg. $d_1 = x_1 - y_1, \dots, d_n = x_n - y_n$. Once you have those values, treat it as a "Estimating $\bar{X}_n$ when σ^2 is/is not known". ← $S_d^2 = \frac{\sum d_i^2 - \frac{(\sum d_i)^2}{n}}{n-1}$ Use $\bar{d}$ as $\bar{x}_n$ ↳ avg of ds

Single Sample: estimating a proportion.
↳ $X \sim \text{Binomial}(n, p)$, want to estimate p. X = # successes, n = # repetitions
Let $\hat{P}$ be estimator of p. Then, a $100(1-\alpha)\%$ CI for p is $\hat{P} \pm Z_{\alpha/2}\sqrt{\hat{P}\hat{Q}/n}$
error will not exceed e when $n = \left(\frac{Z_{\alpha/2}}{e}\right)^2 \hat{p}\hat{q}$, $n \in \mathbb{N}$, note: $\max n = \frac{1}{4}\left(\frac{Z_{\alpha/2}}{e}\right)^2$, $\max n \in \mathbb{N}$

Estimating difference between 2 proportions: a point estimate of the difference $P_1 - P_2$ is $\hat{P}_1 - \hat{P}_2$
↳ if $P_1 - P_2 > 0 \Rightarrow P_1 > P_2$, if $P_1 - P_2 < 0 \Rightarrow P_1 < P_2$, if $0 \in$ CI for $P_1 - P_2$ assume $P_1 = P_2$
a $100(1-\alpha)\%$ CI for $p_1 - p_2$ is $\hat{P}_1 - \hat{P}_2 \pm Z_{\alpha/2}\sqrt{\frac{\hat{P}_1\hat{Q}_1}{n_1} + \frac{\hat{P}_2\hat{Q}_2}{n_2}}$

Estimating Variance (σ^2): use $X^2_{n-1} \sim \frac{(n-1)S^2}{\sigma^2}$, a $100(1-\alpha)\%$ CI for σ^2 is $\left(\frac{(n-1)S^2}{\chi^2_{n-1}((\alpha/2))}, \frac{(n-1)S^2}{\chi^2_{n-1}(1-(\alpha/2))}\right)$

↑ Chi-squared

MAT2377D (Ch6)

Null Hypothesis: H_0. Accept if true, reject if false. H_0 is hypothesis about a parameter

P-Value: smallest level at which H_0 would be rejected. Decision is based on comparing P-Value to α

↳ P-Value $\leq \alpha$ → reject H_0 at level α, P-Value $> \alpha$ → do not reject H_0 at level α

Procedure for Hypothesis Testing:

1- Find H_0 (initial assumption) eg. $H_0: \theta = \theta_0$, where θ is parameter of interest

2- Find alternative hypothesis H_1. H_1 is contradictory to H_0, and is one of the following:

- $H_1: \theta > \theta_0$ ← $P(X \geq x) = p$
- $H_1: \theta < \theta_0$ ← $P(X \leq x) = p$
- $H_1: \theta \neq \theta_0$ ← $P = 2 \cdot P(X \geq x)$ or $P = 2 \cdot P(X \leq x)$

↳ P-value: if $p < 0.05$, reject H_0

3- Find Test Statistic, which is a function of sample data used in decision

4- Decision: reject or not reject H_0 based on rejection region method or p-value

reject if $Z > Z_\alpha$

H_0	Value of Test Statistic	H_1	Critical Region
$\mu = \mu_0$	$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$; σ known	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu = \mu_0$	$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$; $v = n-1$, σ unknown	$\mu < \mu_0$ $\mu > \mu_0$ $\mu \neq \mu_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$; σ_1 and σ_2 known	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$z < -z_\alpha$ $z > z_\alpha$ $z < -z_{\alpha/2}$ or $z > z_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{s_p\sqrt{1/n_1 + 1/n_2}}$; $v = n_1 + n_2 - 2$, $\sigma_1 = \sigma_2$ but unknown, $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$
$\mu_1 - \mu_2 = d_0$	$t' = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$; $v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$, $\sigma_1 \neq \sigma_2$ and unknown	$\mu_1 - \mu_2 < d_0$ $\mu_1 - \mu_2 > d_0$ $\mu_1 - \mu_2 \neq d_0$	$t' < -t_\alpha$ $t' > t_\alpha$ $t' < -t_{\alpha/2}$ or $t' > t_{\alpha/2}$
$\mu_D = d_0$ paired observations	$t = \frac{\bar{d} - d_0}{s_d/\sqrt{n}}$; $v = n-1$	$\mu_D < d_0$ $\mu_D > d_0$ $\mu_D \neq d_0$	$t < -t_\alpha$ $t > t_\alpha$ $t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

Tests Concerning means: ← (see table)

Test on Proportion (single sample): best point estimate for population proportion $\hat{p} = \frac{X}{n}$, $X \sim$ binomial. w/ increasing n, X approaches Normal, so we have

$$Z = \frac{X - np_0}{\sqrt{np_0q_0}} \quad \text{or} \quad Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0q_0}{n}}}$$

↳ testing $H_0: p = p_0$

Other steps are same as in table above (but replace μ w/ p) (2nd row)

Test on Proportion (two-sample):

In this case, $H_0: p_1 - p_2 = 0$, and test statistic is $Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}}$ where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ note: $\hat{p} = \frac{x}{n}$

here, $\hat{p}_1$ would be $\frac{x_1}{n_1}$, $\hat{p}_2$ would be $\frac{x_2}{n_2}$

Once you have Z, compare to Z_α or $Z_{\alpha/2}$ (or t)

if $H_1: p > p_0$, reject if $Z > Z_\alpha$

if $H_1: p < p_0$, reject if $Z < -Z_\alpha$

if $H_1: p \neq p_0$, reject if $Z < -Z_{\alpha/2}$ or $Z > Z_{\alpha/2}$

MAT 2377 D (Ch 7)

Linear Regression: relationship between variables. { Response aka Dependent var: Y / Regressor aka Independent var: X

↳ **Simple Regression**: one regressor (X) to explain Y: $y = \beta_0 + \beta_1 x + \epsilon$

↳ **Multiple Regression**: many regressors to explain Y: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \epsilon$

Simple Linear Regression (SLR): best fitted regression line is $\hat{y} = b_0 + b_1 x_1$ where ↓

$$\text{Slope} = b_1 = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = \frac{(n)(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{(n)(\sum x_i^2) - (\sum x_i)^2} = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} \rightarrow = \frac{S_{xy}}{S_{xx}}$$

$$\text{Intercept} = b_0 = \bar{y} - b_1 \bar{x}$$

b_1 & b_0 are unbiased estimators for β_1 & β_0, so $\mu_{b_1} = \beta_1$, $\mu_{b_2} = \beta_2$

Variance of b_1 & b_0 are $\sigma_{b_1}^2 = \frac{\sigma^2}{\sum (x_i - \bar{x})^2}$, $\sigma_{b_0}^2 = \frac{\sum (x_i^2)}{n \sum [(x_i - \bar{x})^2]}$

Unbiased estimate of σ^2 is $s^2 = \frac{SSE}{n-2} = \frac{\sum (y_i - \hat{y}_i)^2}{n-2} = \frac{S_{yy} - b_1 S_{xy}}{n-2}$, $S_{xx} = \sum [(x_i - \bar{x})^2]$, $S_{yy} = \sum [(y_i - \bar{y})^2]$, $S_{xy} = \sum [(x_i - \bar{x})(y_i - \bar{y})]$

A $100(1-\alpha)\%$ CI for β_1: $b_1 - (t_{\alpha/2}) \frac{s}{\sqrt{S_{xx}}} < \beta_1 < b_1 + (t_{\alpha/2}) \frac{s}{\sqrt{S_{xx}}}$, $t_{\alpha/2}$ is value of T-distr. w/ $n-2$ d.f.

(b_1 ↑ β_1; b_1 ↑ β_1)

Hypothesis testing on slope: To test $H_0: \beta_1 = \beta_{10}$, → $t = \frac{b_1 - \beta_{10}}{s / \sqrt{S_{xx}}}$ ← use t-distr w/ $n-2$ d.f to find critical region or p-value

MAT 2377 D (Ch 8) Control Charts for μ

Probability that $X \in [\mu - 3\sigma, \mu + 3\sigma]$ is 0.9973

[Control Chart] ↓

if we calculate sample average $\bar{x}_n$, it has 0.9973 probably to fall in $\left[\mu - 3\frac{\sigma}{\sqrt{n}}, \mu + 3\frac{\sigma}{\sqrt{n}}\right]$

If μ, σ are unknown: take k samples of n observations each & calculate averages

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \ldots + \bar{x}_k}{k} \qquad \bar{\bar{\sigma}} = \frac{\bar{\sigma}_1 + \ldots + \bar{\sigma}_k}{k}$$

use $\bar{\bar{x}}$ & $\bar{\bar{\sigma}}$ to form the new control chart: $\left[\bar{\bar{x}} - 3\frac{\bar{\bar{\sigma}}}{c\sqrt{n}}, \bar{\bar{x}} + 3\frac{\bar{\bar{\sigma}}}{c\sqrt{n}}\right]$

note: pick c based on n:

n	2	3	4	5	6	7	8	9	10
c	0.5642	0.7236	0.7979	0.8407	0.8686	0.8882	0.9027	0.9139	0.9227

Control Chart for Standard Deviation (σ):

Expected value of sample standard deviation for sample of size n is $c\sigma \leftarrow E(\sigma_n) = c\sigma$,

while variance is $\left(2(n-1) - 2nc^2\right)\frac{\sigma}{\sqrt{2n}} = \sigma^2$

So control charts are $[B_1\sigma, B_2\sigma]$ where $B_2 = c + \frac{3}{\sqrt{2n}}\left(2(n-1) - 2nc^2\right)$, $B_1 = c - \frac{3}{\sqrt{2n}}\left(2(n-1) - 2nc^2\right)$

To save time, values for B_1, B_2 are given:

n	2	3	4	5	6	7	8	9	10
B_1	0	0	0	0	0.026	0.105	0.167	0.219	0.262
B_2	1.843	1.858	1.808	1.756	1.711	1.672	1.638	1.609	1.584

If σ is unknown: take k samples of n observations each,

$\bar{\bar{\sigma}} = \frac{\bar{\sigma}_1 + \ldots + \bar{\sigma}_k}{k}$, estimate $\sigma \cong \frac{\bar{\bar{\sigma}}}{c}$ $\Rightarrow$ and then use $[B_3\sigma, B_4\sigma]$

where $B_4 = 1 + \frac{3}{\sqrt{2n}c}\left(2(n-1) - 2nc^2\right)$ & $B_3 = 1 - \frac{3}{\sqrt{2n}c}\left(2(n-1) - 2nc^2\right)$

Control Chart for % defective (p): take k subgroups of n items, count #defectives = D

each subgroup and proportion of defectives $p = \frac{D}{n}$ eg. for $k=2, n=50$

subgroup	D	p
1	1	0.02
2	5	0.10

$\Big\} \bar{p} = \frac{6}{100}$

↳ $\bar{p} = \frac{\text{total \#defectives}}{kn}$ and control chart is: $[LCL^*, UCL]$

$$UCL = \bar{p} + 3\frac{\sqrt{\bar{p}(1-\bar{p})}}{\sqrt{n}} \qquad LCL^* = \bar{p} - 3\frac{\sqrt{\bar{p}(1-\bar{p})}}{\sqrt{n}}$$

* Set $LCL = 0$ if $LCL < 0$